8.3 LE ROUTAGE DANS LE RÉSEAU IP

8.3.1 L'adressage d'interface

Supposons le réseau simplifié de la figure 8.27, peu importe le protocole mis en œuvre sur le lien reliant les passerelles d'accès (routeurs IP). Comment la couche IP peut-elle déterminer l'interface de sortie par rapport à une adresse IP destination alors que la couche IP ignore la technologie sous-jacente ?

En application du principe d'indépendance des couches, le point d'accès au réseau physique ne peut être connu de la couche IP que par une adresse IP. Les liaisons entre les différentes passerelles sont considérées, vu d'IP, comme constituant un réseau ; de ce fait chaque extrémité d'une liaison possède une adresse IP. Dans ces conditions, l'algorithme d'acheminement recherche sur quel réseau (de liaisons) est situé le saut suivant. Cette technique d'identification de l'interface d'accès au réseau physique est dite adressage d'interface ou **adressage de LS** (Liaison spécialisée). Cette méthode garantit l'indépendance des couches. En effet, le routage se réalise d'adresse IP destination à adresse IP d'interface (*Next hop*).

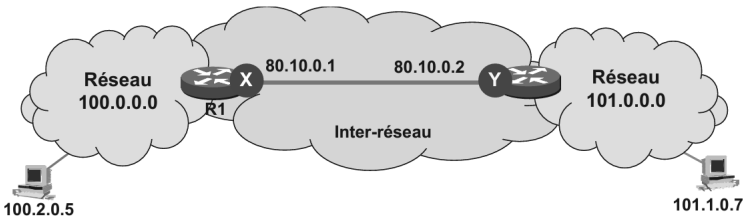

Figure 8.27 Adressage du réseau physique.

Les liens inter-routeurs forment ainsi le réseau logique IP. Une adresse IP est attribuée à chaque extrémité. Pour comprendre le mécanisme de routage, la figure 8.28 fournit un exemple de configuration d'un routeur[15]. Notons que la route à prendre est désignée par l'adresse distante du lien (*Next hop*), ce qui correspond à l'adresse du point à atteindre sur le réseau de liens.

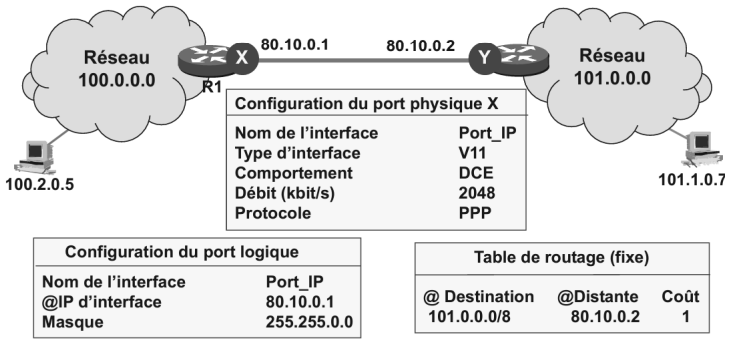

Figure 8.28 Exemple de configuration d'un routeur IP.

15. Le mode de configuration d'un routeur est spécifique à chaque constructeur. Aussi, les exemples ci-après ne constituent qu'une illustration particulière.

8.3.2 Le concept d'interface non numérotée

L'attribution d'adresses d'interface est consommatrice d'adresses, aussi le RFC 1812 a-t-il autorisé le routage sur interface dite non numérotée (*Unnumbered IP*). Un exemple de configuration simplifiée est donné par la figure 8.29. Cette approche viole la règle d'indépendance des couches. Aussi, le RFC 1812 précise que les deux passerelles connectées par une ligne point à point non numérotée ne sont pas à considérer comme deux passerelles distinctes mais comme deux demi-passerelles constituant une seule passerelle virtuelle.

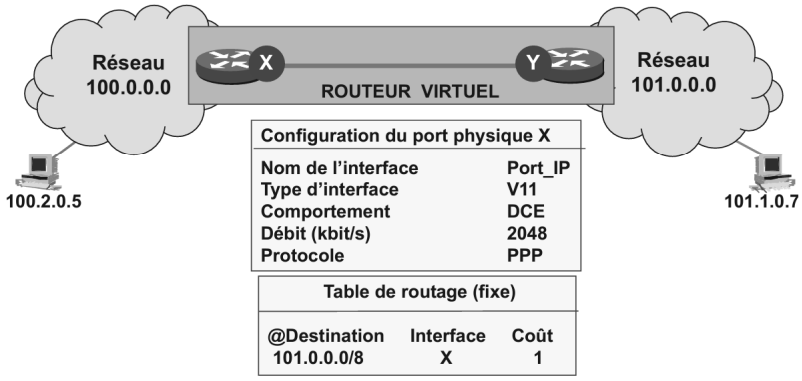

Figure 8.29 Routage du RFC 1812.

Les limitations de l'adressage, dues essentiellement à la structure à plat de l'adressage IP, conjuguées au succès d'Internet ont contribué à la pénurie d'adresses. Plusieurs solutions ont été imaginées pour pallier la pénurie. Celle retenue, IPv6, porte l'espace d'adressage de 32 à 128 bits.

8.4 L'ADRESSAGE DANS IPv6 (RFC 4291)

8.4.1 Généralités

IPng (*Next generation*) ou IPv6 répond au besoin d'évolution de la communauté Internet et comble les faiblesses d'IPv4 dont la plus connue concerne l'espace d'adressage. IPv4 met en place un adressage à plat (Net_ID) ce qui a conduit à l'explosion des tables de routage (certains routeurs Internet ont plusieurs dizaines de milliers d'entrées dans leur table de routage). En autorisant l'agrégation d'adresses de réseaux contiguës en un seul préfixe réseau, en organisant une affectation géographique des adresses et en faisant disparaître la notion de classes d'adresses, le CIDR a en partie répondu à ce problème. La seconde faiblesse concerne la pénurie d'adresses[16], l'utilisation d'un adressage privé associé à la translation d'adresses (NAT) résout partiellement ce problème mais pénalise fortement les performances. Aussi, l'adressage dans IPv6 a-t-il été étendu à 128 bits, ce qui correspond à plusieurs milliers d'adresses au m^2 de surface terrestre ($6{,}65 \cdot 10^{23}$@/m^2).

Dans un système de réseaux interconnectés, seul un adressage hiérarchique permet l'allégement des tables de routage, chaque routeur ne traitant que la partie de l'adresse correspondant à son domaine. Cependant, dans une communauté aussi vaste que celle d'Internet, l'adressage

16. Depuis février 2011, l'IANA ne dispose plus d'adresse IPv4.

hiérarchique devient vite sans signification, aussi entre l'adressage à plat non significatif d'IPv4 et l'adressage hiérarchique géographique, tel que celui d'X.121, un compromis a été réalisé. L'adressage IPv6 comporte quatre champs (figure 8.30).

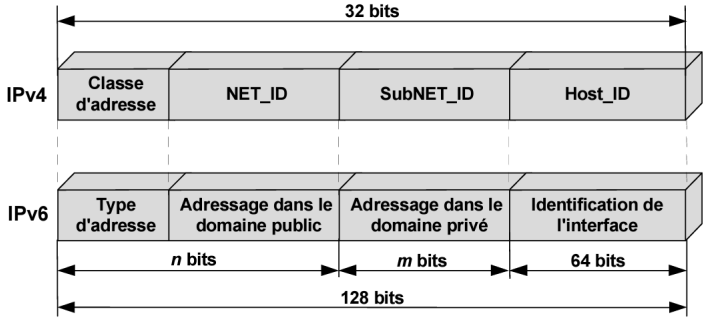

Figure 8.30 Principe de l'adressage IPv6.

Le premier, sur *n* bits, est une agrégation hiérarchique de préfixes décrivant la connectivité du site, ce champ est désigné sous le terme topologie publique, il est attribué par l'opérateur auquel on est raccordé. L'affectation d'adresses par l'opérateur change radicalement la nature de l'adressage, on n'est plus propriétaire de ses adresses ; si on change d'opérateur, on change d'adresse !

Le second, sur *m* bits, décrit la topologie locale du site, enfin le dernier, sur 64 bits, identifie de manière unique une interface. Cet adressage est dénommé adressage agrégé ou *Aggregatable Global Unicast Address Format*.

8.4.2 La notation IPv6

Une notation hexadécimale, sur 16 bits séparés par deux points « : », remplace la notation décimale pointée d'IPv4. L'adresse passe de 32 à 128 bits, huit mots de 16 bits. Ainsi, une adresse IPv6 s'écrit :

FE0C:DA98:0:0:0:0:5645:376E

La notation peut être simplifiée en remplaçant une succession de 0 par « :: », l'abréviation « :: » ne pouvant être utilisée qu'une seule fois. Ainsi, l'adresse précédente devient :

FE0C:DA98::5645:376E

IPv6 adopte une notation similaire à celle du CIDR, le champ préfixe étant désigné par un nombre représentant la longueur en bits du préfixe, l'écriture est donc de la forme : @IPv6/ longueur du préfixe en bits, soit par exemple :

FE0C:DA98/32
FE0C:DA98:0:0/64
FE0C:DA98::/64

Il est encore possible de simplifier l'écriture en supprimant les zéros en tête d'un bloc de 16 bits, ainsi l'adresse « FE0C:DA90:0:0:0:0:0645:3760 » peut s'écrire :

FE0C:DA90::645:3760 (écriture correcte),
mais pas FE0C:DA9::645:376 (écriture incorrecte)

Une adresse IP peut être utilisée comme URL (*Uniform Ressource Locators*), par exemple en IPv4 : http://80.12.4.212:8080 où :8080 désigne l'application distante (port). En IPv6, on obtiendrait une URL du type : http:// FE0C:DA98::5645:3763:8080 ce qui introduit une ambiguité, 8080 est le port destination où le dernier champ de l'adresse ? Aussi, le RFC 2732 propose d'écrire l'adresse IPv6 entre « [ ] », dans notre exemple cela donne :

http://[FE0C:DA98::5645:3763]:8080

8.4.3 Les types d'adresse

Très pénalisante en termes de performance réseau, la notion de *broadcast* disparaît. Elle est remplacée par une généralisation des adresses *multicast*.

IPv6 distingue trois types d'adresse :

- les adresses ***unicast*** (*one-to-one*) : une adresse *unicast* désigne une interface, elle peut être utilisée pour identifier un groupe d'interfaces lorsque ces interfaces constituent une agrégation de liens et qu'ils doivent être vus comme une seule interface ;
- les adresses ***multicast*** (*one-to-any*) : ces adresses désignent un ensemble d'interfaces dont la localisation n'est pas nécessairement sur le même réseau physique. Un datagramme adressé à une adresse *multicast* est acheminé à toutes les interfaces du groupe ;
- les adresses ***anycast*** (*one-to-nearest*) : ces adresses introduites par IPv6 correspondent à une restriction des adresses de *multicast*. Elles désignent un ensemble d'interfaces partageant un même préfixe réseau. Cependant, lorsqu'un datagramme est adressé à une adresse *anycast*, il n'est délivré qu'à une seule interface du groupe, celle dont la métrique, au sens routage du terme, est la plus proche du nœud source[17].

8.4.4 Le plan d'adressage

L'identifiant d'interface

L'identifiant d'interface dans IPv6 correspond à la notion d'Host_ID d'IPv4. Afin de faciliter les opérations d'autoconfiguration, de disposer d'un identifiant unique au niveau mondial et de répondre aux besoins nouveaux des réseaux domotiques (réseaux IEEE 1394, *Firewire*), l'IEEE a étendu la numérotation des interfaces à 64 bits (*EUI-64*, *End-User Interface*). Pour les interfaces non dotées de cet identifiant, celui-ci est construit à partir de l'adresse MAC IEEE de 48 bits dite MAC-48 (figure 8.31).

Les bits U et G ont la même signification que les bits U/L des adresses MAC-48 de l'IEEE :

- **U** (Universel), à 1 il indique l'universalité de l'identifiant d'interface, à zéro il indique qu'il s'agit d'un identifiant géré localement par l'administrateur du réseau, la signification de ce bit est inversée par rapport à celle du bit U/L de l'adressage MAC-48. Pour convertir une adresse MAC-48 en une adresse EUI-64, il convient donc d'inverser (complémenter) le bit U/L.
- **G** (Global), à zéro, l'identifiant désigne une interface unique (*unicast*), à 1 il indique qu'il s'agit d'un identifiant de diffusion et qu'il adresse plusieurs interfaces (*multicast*).

17. Ces adresses qui ne désignent précisément personne sont parfois appelées : « adresse à la cantonnade ».

La représentation diffère de celle de l'adresse MAC-48 qui représente les bits dans l'ordre d'émission alors que l'adresse EUI-64 est requise que le I64 positionne les bits de poids forts devant, cependant l'émission des bits reste identique. Les bits de poids faibles sont toujours émis en premier.

L'adresse IPv6 d'un Host est constituée de la concaténation d'un préfixe et de l'identifiant d'interface (EUI-64, figure 8.30). Cette méthode de détermination du champ « Host-ID » pour reprendre le langage d'IPv4 autorise un mécanisme d'autoconfiguration des hôtes. Une machine (passerelle par défaut en général) diffuse périodiquement le préfixe utilisé localement sur le réseau. Il suffit à une machine sans adresse d'ajouter à ce préfixe son identifiant d'interface et de vérifier l'unicité de cette adresse.

L'adresse IPv6 d'un host étant liée à celle de son interface physique (figure 8.30), en cas de changement de carte réseau l'adresse IP de la machine change, ce qui, pour des serveurs et en particulier le DNS, peut constituer un handicap. Aussi, pour les machines devant conserver un adressage fixe, il est recommandé d'attribuer manuellement un identifiant d'interface.

Cette méthode de détermination du champ « Host_ID » pour reprendre le langage IPv4 autorise un mécanisme d'autoconfiguration des hôtes. Une machine, la passerelle par défaut en général, diffuse périodiquement le préfixe utilisé localement sur le réseau ; il suffit alors à une machine sans adresse d'ajouter à ce préfixe son identifiant d'interface et de vérifier l'unicité de cette adresse construite.

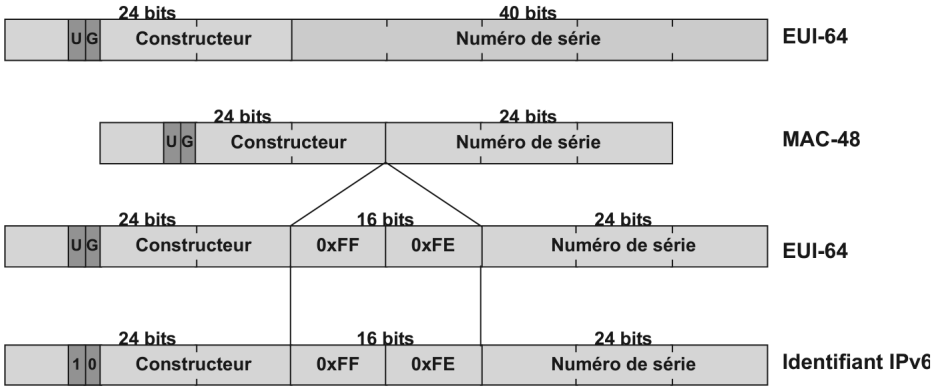

Figure 8.31 Construction de l'identifiant IPv6 à partir de l'adresse MAC.

Identifiant d'interface et vie privée (RFC 3041)

Toute adresse IPv6 quel que soit le préfixe identifie un hôte.

Il est donc aisé de déterminer tous les déplacements de l'utilisateur de cette machine. Aussi, le RFC 3041 a-t-il proposé un mécanisme d'affectation aléatoire d'identifiant d'interface (identifiant privé, bit U à « 0 »), l'adresse est dite temporaire.

Périodiquement, cet identifiant est renouvelé. Les communications établies se poursuivent sur l'ancienne adresse (adresse dite à l'état déprécié), les nouvelles communications utiliseront la nouvelle adresse (adresse dite à l'état préféré).

Le RFC 3972 introduit un mécanisme de génération à partir de la clé publique de la machine (CGA, *Crytographic Generated Addresses*) sécurisant ainsi les mécanismes associés comme la découverte des voisins (procotole remplaçant l'ARP traditionnel d'IPv4).

L'adressage unicast

➤ Adressage global unicast

L'adressage agrégé est un adressage hiérarchique à trois niveaux (figure 8.32).

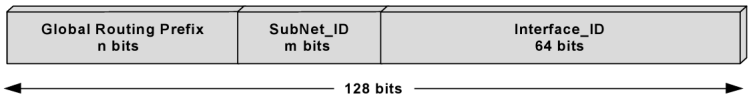

Figure 8.32 Structure de l'adressage unicast global (RFC 4291).

Seules routables sur Internet, ces adresses sont définies comme appartenant au sous-réseau 2000::/3. L'adressage agrégé comporte trois champs :

- ❏ un préfixe global (*Global Routing Prefix*) sur *n* bits définissant la topologie publique, attribuée par l'opérateur, elle détermine le routage dans Internet ;
- ❏ les *m* bits suivants décrivent la topologie privée, ce champ à la même signification que le champ SubNet_ID d'IPv4, il est parfois désigné sous l'appellation de **SLA** (*Site Level Aggregator*). L'identification des sous-réseaux est de la responsabilité de l'administrateur local ;
- ❏ enfin, le dernier champ sur 64 bits identifie l'interface raccordée au réseau.

Transitoirement, avant l'ouverture du registre officiel, des adresses de préfixe 2000::/16 ont été attribuées. Elles ne sont plus valides aujourd'hui.

Le tableau de la figure 8.33 répertorie les adresses valides et routables sur l'Internet. On y distingue :

- ❏ les adresses 2001::/16 et suivantes qui sont ouvertes à la réservation depuis 2001 ;
- ❏ les adresses 6to4 (2002::/16) qui appartiennent au plan d'adressage de transition permettant l'acheminement du trafic IPv6 via un ou plusieurs réseaux IPv4.

À l'origine, l'IANA allouait aux RIR (*Regional Internet Registries*) des blocs de 23 bits (exemple : 2001:0600::/23) ; ces mêmes RIR distribuent aux LIR (*Local Internet Registries*) des blocs de 32 bits qui affectent des blocs de 48 bits aux utilisateurs finaux (RFC 3177). Cette rigidité d'affectation a été abolie par le RFC 6177 (figure 8.34).

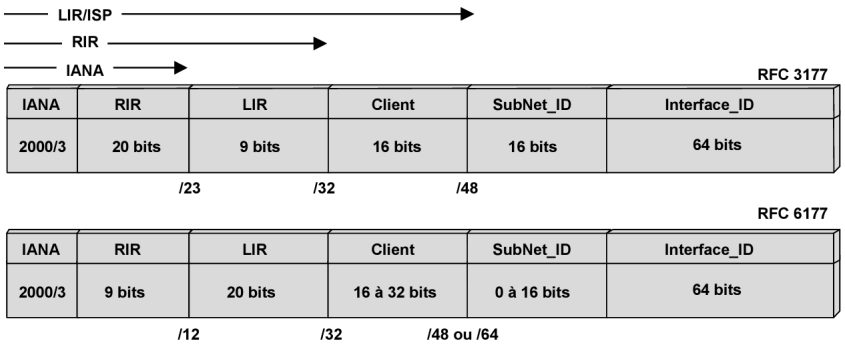

Figure 8.34 Allocation d'adresses.

Préfixe global	Assignation	Date
2001:0000::/23	IANA	01 Juillet 1999
2001:0200::/23	APNIC	01 Juillet 1999
2001:0400::/23	ARIN	01 Juillet 1999
2001:0600::/23	RIPE NCC	01 Juillet 1999
2001:0800::/23	RIPE NCC	01 Mai 2002
2001:0A00::/23	RIPE NCC	02 Novembre 2002
2001:0C00::/23	APNIC	01 Mai 2002
2001:0E00::/23	APNIC	01 Janvier 2003
2001:1200::/23	LACNIC	01 Novembre 2002
2001:1400::/23	RIPE NCC	01 Février 2003
2001:1600::/23	RIPE NCC	01 Juillet 2003
2001:1800::/23	ARIN	01 Avril 2003
2001:1A00::/23	RIPE NCC	01 Janvier 2004
2001:1C00::/22	RIPE NCC	01 Mai 2004
2001:2000::/20	RIPE NCC	01 Mai 2004
2001:3000::/21	RIPE NCC	01 Mai 2004
2001:3800::/22	RIPE NCC	01 Mai 2004
2001:3C00::/22	RESERVED	11 Juin 2004
2001:4000::/23	RIPE NCC	11 Juin 2004
2001:4200::/23	AfriNIC	01 Juin 2004
2001:4400::/23	APNIC	11 Juin 2004
2001:4600::/23	RIPE NCC	17 Août 2004
2001:4800::/23	ARIN	24 Août 2004
2001:4A00::/23	RIPE NCC	15 Octobre 2004
2001:4C00::/23	RIPE NCC	17 Décembre 2004
2001:5000::/20	RIPE NCC	10 Septembre 2004
2001:8000::/19	APNIC	30 Novembre 2004
2001:A000::/20	APNIC	30 Novembre 2004
2001:B000::/20	APNIC	08 Mars 2006
2002:0000::/16	6to4	01 Février 2001
2003:0000::/18	RIPE NCC	12 Janvier 2005
2400:0000::/12	APNIC	03 Octobre 2006
2600:0000::/12	ARIN	03 Octobre 2006
2610:0000::/23	ARIN	17 Novembre 2005
2620:0000::/23	ARIN	12 Septembre 2006
2800:0000::/12	LACNIC	03 Octobre 2006
2A00:0000::/12	RIPE NCC	03 Octobre 2006
2C00:0000::/12	AfriNIC	03 Octobre 2006
2D00:0000::/8	IANA	01 Juillet 1999
2E00:0000::/7	IANA	01 Juillet 1999
3000:0000::/4	IANA	01 Juillet 1999
3ffe::/16	IANA	Avril 2008
5f00::/8	IANA	Avril 2008

Figure 8.33 Synthèse des préfixes alloués au 30 mai 2012.
Source : http://www.iana.org/assignments/ipv6-unicast-address-assignments

➤ Adressage de lien local (*Unique Local Address*)

Initialement[18] des adresses de site local (FEC0::/10, figure 8.34) assimilables aux adresses privées d'IPv4 (RFC 1918) avaient été imaginées. À l'instar d'IPv4, ces adresses n'avaient qu'une portée

18. L'utilisation de cet adressage, aujourd'hui retiré du standard, est fortement déconseillée.

limitée au réseau privé, elles ne pouvaient être utilisées pour se connecter à Internet. Le champ **SLA_id** (*Site Level Aggregator*) correspondant au champ Subnet_ID d'IPv4 permettant à l'administrateur local de hiérarchiser son réseau. N'utilisant aucun préfixe d'identification ces adresses posaient de nombreux problèmes (RFC 3879).

Ces adresses ont été remplacées par des adresses dites uniques locales (**ULA**, *Unique Local Address*). Ces adresses n'ont évidemment qu'une portée locale mais peuvent être routées dans le domaine privé. La figure 8.35 décrit ce nouveau format d'adresses.

- ❏ Le champ *Préfix* sur 7 bits identifie le format de l'adresse (FC00::/7) ;
- ❏ Le bit suivant (*L*) actuellement toujours à 1, indique que le préfixe a été attribué localement ;
- ❏ Le champ *Gobal ID* (*Globally Unique Prefix*) sur 40 bits permet l'identification du site (nombre pseudo-aléatoire) ;
- ❏ Le champ *SubNet_ ID* sur 16 bits, identifie un sous-réseau ;
- ❏ Enfin le champ *Interface ID* (64 bits).

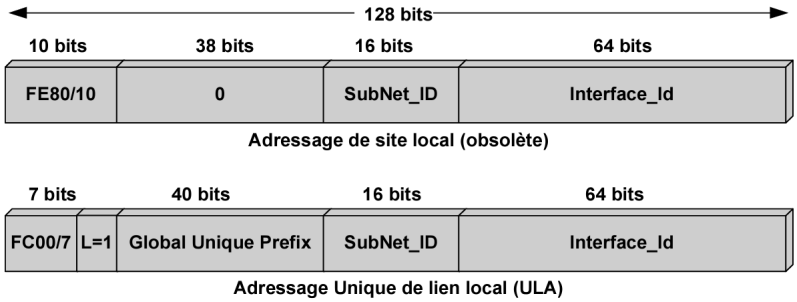

Figure 8.35 Ex-adressage de site local (adressage privé).

➤ Adressage de lien local

Les adresses de lien local identifient des interfaces de connectivité directe, sans passerelle intermédiaire, ce sont par exemple les machines reliées directement entre elles (adressage des LS d'IPv4), les machines d'un même brin de réseau local (Ethernet ou autre). Ces adresses sont définies automatiquement lors de l'initialisation de l'interface. Elles correspondent à la concaténation de l'identifiant d'interface et du préfixe FE80::/64 (figure 8.36).

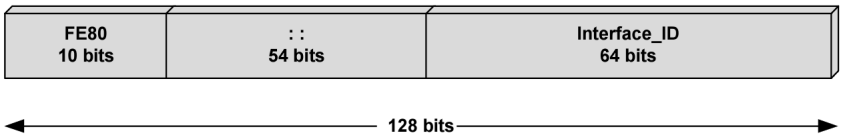

Figure 8.36 Adressage de lien local.

Les adresses de lien local ne sont pas routables, ni sur Internet, ni dans le domaine privé elles ne sont utilisées que par certains protocoles notamment lors de la configuration d'adresse globale et par les protocoles de découverte de voisins et de routeurs. En principe, leur utilisation est à proscrire comme adresse d'hôte dans les applications traditionnelles.

➤ Adresses spécifiques

Comme dans IPv4 certaines adresses ont une signification particulière :

- ❏ L'adresse dite indéterminée (*Unspecified address*), correspondant à l'adresse 0.0.0.0 d'IPv4, elle désigne une interface en cours d'initialisation. Cette adresse 0:0:0:0:0:0:0:0 ou « ::/128 » ne doit jamais être attribuée à une interface.
- ❏ L'adresse de bouclage (*Loopback address*) correspond à l'adresse 127.0.0.1 d'IPv4, elle vaut 0:0:0:0:0:0:0:1 ou encore ::1/128.

L'adressage multicast

Afin d'éviter l'usage intempestif de *broadcast* qui pénalise les performances, IPv6 a généralisé la notion de *multicast* en les spécialisant et en définissant différents niveaux de diffusion. Une adresse *multicast* désigne un ensemble de nœuds, elle est dite aussi adresse sur abonnement.

L'adresse *multicast* comporte quatre champs (figure 8.37) :

- ❏ Le premier identifie une adresse de *multicast* (préfixe FF00::/8).
- ❏ Le deuxième, le champ sur 4 bits contient les indicateurs « **0 R P T** » :
  - le bit **T** (*Transient*) à « 0 », indique que l'adresse est permanente (adresse affectée par l'IANA), à « 1 » que l'adresse est temporaire (transitoire) ;
  - le bit **P** indique le format de l'adressage : à « 0 » il signifie que le format d'adresse est générique, à « 1 » il indique que l'adresse multicast est dérivée de l'adressage unicast de l'organisation. Le bit « P=1 » implique que le « bit T=1 » (RFC 3306). La longueur du préfixe utilisé est donnée par le champ « **Plen** » (*Prefix Length*). Ce format d'adresse multicast permet à une organisation de disposer de plus de 4 milliards d'adresses multicast routables sur internet.
  - Le bit **R** est réservé, il est positionné à « 0 ».
- ❏ Le champ suivant indique le niveau de diffusion (*scope*). Les différentes valeurs de ce champ sont :
  - 0, réservé ;
  - 1, nœud local (*node local scope*) ;
  - 2, lien local (*link local scope*) ;
  - 5, site local (*subnet local scope*) ;
  - 8, l'organisation locale (*organisation local scope*) ;
  - E, global (*global scope*) ;
  - F, réservé.
- ❏ Enfin le dernier est sur 112 bits

Certaines adresses *multicast* ont été prédéfinies :

- ❏ FF02::1, tous les hôtes d'un même lien ;
- ❏ FF02::2, tous les routeurs du même lien que l'expéditeur ;
- ❏ FF05::5, tous les routeurs du même site que l'expéditeur ;
- ❏ FF05::B, Home agent (mobilité) ;
- ❏ FF05:: 1:2, tous les agents DHCP ;
- ❏ FF05:: 1:3, tous les serveurs DHCP ;
- ❏ FF0*x*:: 1:FF/104, préfixe d'une adresse « *multicast* sollicité ».

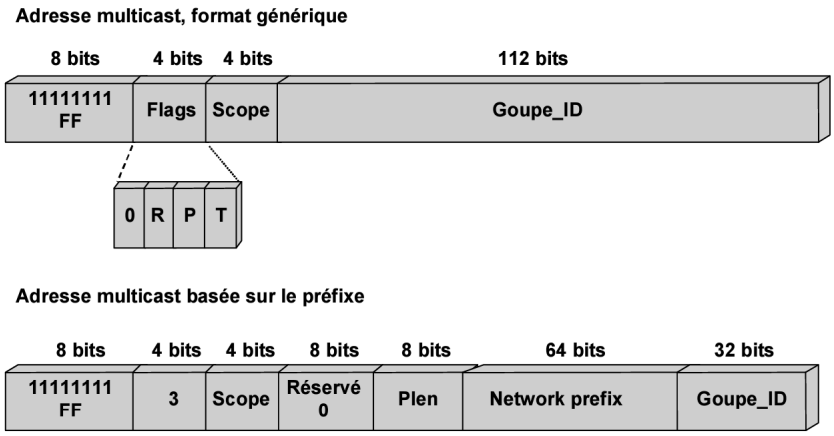

Figure 8.37 Adressage multicast.

De même, la suppression de la notion de *broadcast* conduit à remplacer les adresses physiques IEEE MAC de *broadcast* par des adresses MAC *multicast*. Ces adresses sont construites à partir du préfixe « 33:33 » et des 24 derniers bits de l'adresse IP Multicast. La figure 8.38 fournit un exemple d'adresses *multicast*.

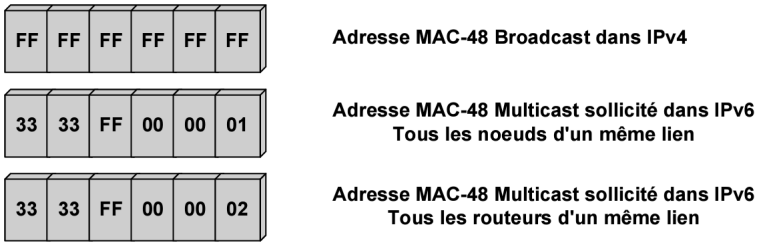

Figure 8.38 Adressage IEE MAC multicast.

À l'initialisation, une machine doit adhérer à deux groupes d'adresses multicast :

- ❏ tous les nœuds du lien « FF02::1 »,
- ❏ le multicast sollicité de la machine, constitué du préfixe « FF02::1:FFxx:xxxx/104 » et des 24 derniers bits de l'adresse IPv6 de la machine.

Soit pour la machine d'adresse IP « 2002:0600:1001:0F02:0200:CFF:FE**04:0506** »

- ❏ tous les nœuds du lien « FF02::1 »,
- ❏ multicast sollicité « FF02:0000:0000:0000:0000:0001:FF**04:0506 »**
- ❏ sur le réseau Ethernet « 33:33:FF: **04:05:06 »**

Les adresses de « multicast sollicité » sont utilisées pour joindre un nœud dont on connaît l'adresse IP, mais pas l'adresse MAC-48.

L'adressage d'anycast

Concrètement, une adresse *anycast* (RFC 4291) résulte de la concaténation d'un préfixe désignant le sous-réseau adressé et d'un suffixe nul (figure 8.39). En principe, actuellement une adresse *anycast* ne peut être attribuée qu'à une passerelle. L'adressage *anycast* désigne plus un service qu'une machine.

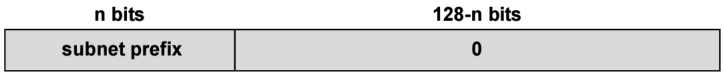

Figure 8.39 Adressage d'anycast.

8.4.5 Migration IPv4 vers IPv6

Même si la fin de la disponibilité d'adresses IPv4 devrait accélérer le processus[19], la migration des systèmes vers IPv6 ne se fera que très progressivement. La raison en est simple pourquoi changer un système qui fonctionne ? Cependant, si elle tarde dans les éléments de réseaux des entreprises, elle est très avancée chez les opérateurs et dans tous les nouveaux systèmes, en particulier pour les accès Internet de la téléphonie mobile. Pour faciliter cette transition plusieurs solutions sont avancées, elles reposent sur les principes suivants, utilisés seuls ou en combinaison :

- ❏ La technique de double pile (*Dual-stack*), IPv4 et IPv6 cohabitent sur le même nœud.
- ❏ Les techniques de tunnel qui encapsulent le datagramme d'origine dans le protocole destination.
- ❏ Les techniques de translation qui adaptent le datagramme au protocole du réseau cible (conversion de protocole).

Principe des différentes techniques

➤ Principe de la technique de la double pile

Cette solution dite *dual-stack* (*DSTM, Dual Stack Transition Mechanism*), la plus simple a priori, consiste à mettre en œuvre sur chaque nœud du réseau (machines terminales, serveurs, commutateurs, routeur...) les deux piles de protocole. Chaque interface est dotée d'une adresse IPv4 et d'une adresse IPv6. Ainsi, dans la figure 8.40 la machine « *A* » peut communiquer en IPv4 avec la machine « *B* », la machine « *C* » peut aussi communiquer avec « B » en IPv6, quant à la communication entre « *A* » et « *C* » elle est impossible (communications dites *4to4* et *6to6*).

Cette solution permet le déploiement d'un réseau IPv6 sur un réseau IPv4. Cependant, si cette migration a le mérite de la simplicité, celle-ci n'est qu'apparente, elle devient vite complexe dans un grand réseau comme ceux des opérateurs et elle n'apporte aucune réponse au problème de la pénurie d'adresse IPv4.

Cette approche nécessite l'implémentation de 2 piles TCP/IP complète. Il est possible de n'utiliser qu'une couche transport commune en représentant en interne les adresses IPv4 au format IPv6 (adresse IPv4 mappée figure 8.41). Les clients IPv6 sont gérés normalement et les clients IPv4 comme des clients IPv6.

19. Rappelons que depuis le 3 février 2011, l'IANA ne dispose plus d'aucune adresse IPv4. Le RIPE, début 2012, ne disposait plus que d'environ 150 000 adresses.

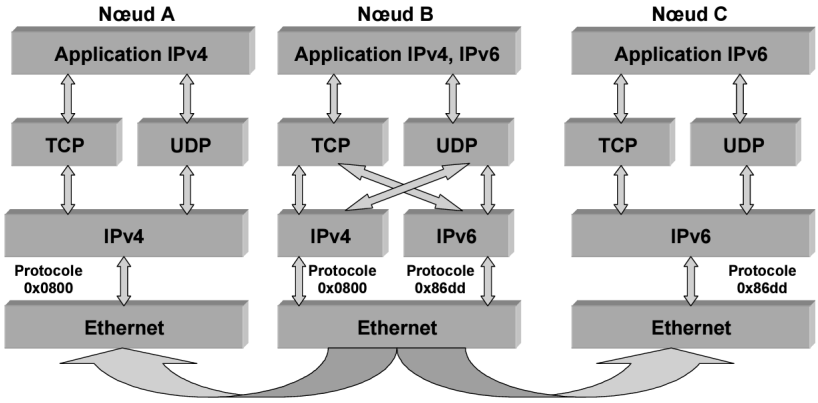

Figure 8.40 Principe de la double pile.

Figure 8.41 Adressage IPv4/IPv6.

➤ Principe de la technique du tunnel

Une alternative au déploiement massif d'un système *dual-stack* consiste à utiliser des tunnels pour le transport IPv6 dans IPv4 (transit de données IPv6 sur un réseau IPv4) ou l'inverse transporter de l'IPv4 sur une infrastructure IPv6 (figure 8.42).

Les tunnels peuvent être statiques (configurés par l'administrateur) ou dynamiques. Cette méthode voit tout son intérêt lors d'une migration d'un réseau IPv4 vers IPv6. La passerelle d'accès au réseau examine le datagramme, si le datagramme d'arrivée correspond au protocole du réseau de transit, le datagramme est acheminé nativement ; si ce n'est pas le cas, il sera encapsulé dans un datagramme du protocole du réseau de transit. Les techniques de tunnels sont appropriées pour assurer des communications *4to4* et *6to4 via 6*.

➤ La translation de protocole

Compte tenu du peu de différence des informations d'en-tête entre IPv6 et IPv4, il est tout à fait concevable, non seulement de mettre en correspondance une adresse IPv4 et une adresse IPv6 (NAT[20], *Network Address Translator*) mais directement d'assurer une conversion de protocole, c'est l'objet de la translation. Les translateurs sont des équipements qui assurent une conversion de protocole d'IPv4 vers IPv6 et vice-versa d'IPv6 vers IPv4. Cette technique illustrée figure 8.43 autorise des communications *4to6* via un réseau IPv6 et *6to4* via un réseau IPv4.

Si cette solution permet de s'affranchir de la double pile de protocole dans les nœuds, elle nécessite l'utilisation d'un DNS ALG (*Application Level Gateway*) qui interfère sur la communication *End to End* et pose notamment des problèmes avec l'utilisation IPSec[21].

---

20. À l'origine le NAT mettait en correspondance une adresse IPv4 publique et une adresse IPv4 privée (NAT44).
21. IPSec ou IP *Secure* (voir § 23.3) est un protocole de sécurité masquant notamment les adresses source et destination.

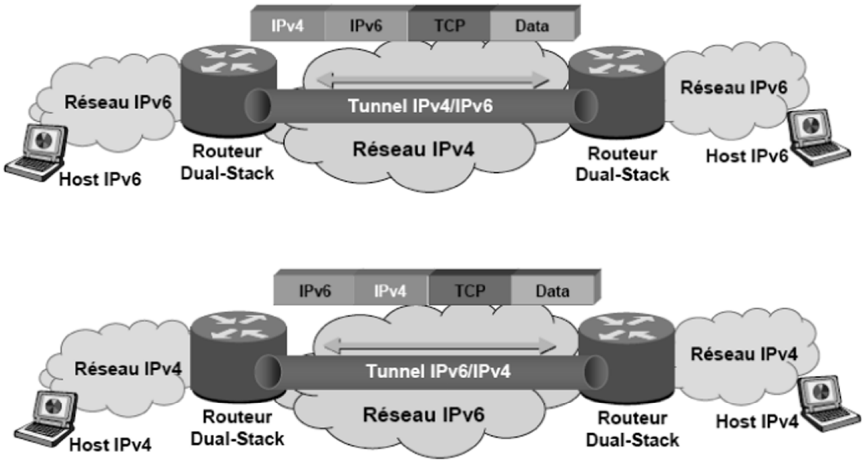

Figure 8.42 Principe des tunnels IPv6/IPv4 et IPv4/IPv6.

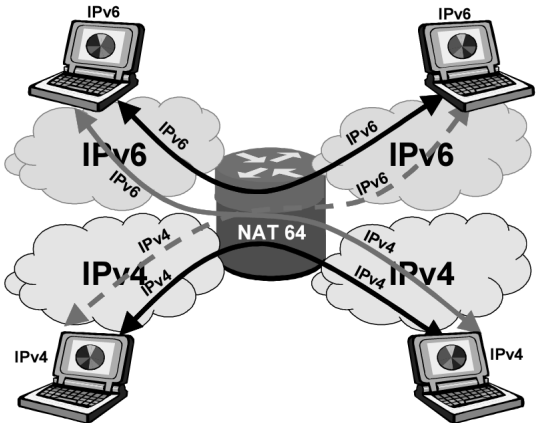

Figure 8.43 Principe de la translation de protocole.

Exemples de mise en œuvre

➤ Les tunnels 6to4 (RFC 3056)

La technique *6to4* offre une connectivité IPv6 au travers d'une infrastructure IPv4 sans configuration explicite de tunnels (figure 8.44). Les tunnels sont initialisés par la passerelle *6to4* (encapsulation des datagrammes IPv6 dans des datagrammes IPv4).

Le routage *6to4* est asymétrique (figure 8.43), le paquet à destination de l'Internet IPv6 est adressé à l'adresse unicast réservée aux passerelles *6to4* « 192.88.99.1 » (RFC 5735 : 192.88.99.0/24). Le trafic retour est directement acheminé vers une passerelle *6to4* d'adresse « 2002 ::/16 ».

Une plage d'adresses spécifiques (préfixe 2002 ::/16) a été réservée aux raccordements de sites IPv6 ne disposant pas de préfixe alloué par le fournisseur d'accès, cependant, le site doit disposer

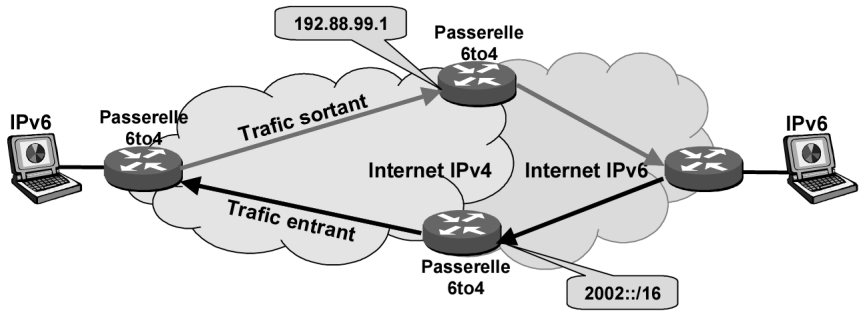

Figure 8.44 Principe du réseau 6to4.

d'une adresse IPv4 publique. Supposons qu'un site ait pour adresse publique IPv4 : 192.0.12.150, la figure 8.45 montre la construction du préfixe IPv6 du site.

Préfixe IPv6 6to4	2002	Adresse publique IPv4 du site			/48
Adresse IPv4 du site	192	0	12	150	
	C0	00	0C	96	
Préfixe IPv6 du site	**2002**	**C000**	**0C96**		/48

Figure 8.45 Détermination du préfixe IPv6 du site.

Longtemps considéré comme la solution mais aujourd'hui, compte tenu de nombreux problèmes en relation avec de la perte de trafic, *6to6* pourrait bien être relégué en « RFC historique ».

➤ Le 6rd (rapid deployment, RFC 5969)

Lors de la transition d'un réseau d'opérateur vers le protocole IPv6, ce n'est pas la migration de son cœur de réseau qui est délicate, mais bien celle de son réseau de collecte qui assure non seulement la collecte du trafic mais aussi garantit la sécurité des accès. Le 6rd est une évolution du 6to4 dédiée au réseau de collecte d'un fournisseur d'accès à Internet. Cette technique a été mise en œuvre dès 2007 par Free. La technique 6rd est représentée figure 8.46.

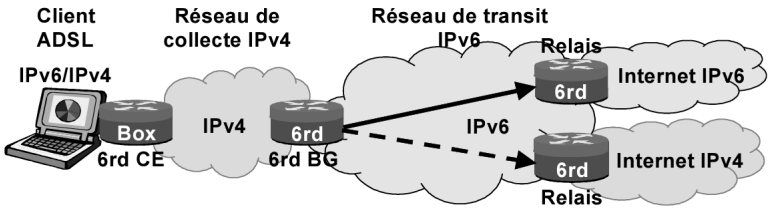

Figure 8.46 Déploiement du 6rd.

Le 6rd comprend deux équipements. Côté résidentiel, la passerelle d'accès (Box ADSL) au réseau de collecte du FAI (Fournisseur d'accès à Internet) qui correspond dans la terminologie 6rd à l'équipement dit « **6rd CE** » (*Customer Equipment*) est dotée d'une double pile (DS) IPv4/IPv6 elle initialise le tunnel IPv4. Côté fournisseur d'accès, l'équipement de bordure (**6rd BR**, *Border Router*) réalise l'interface entre le réseau IPv4 de collecte et le réseau de transit IPv6 de l'opérateur.

Le trafic IPv4/IPv6 du client est encapsulée dans un paquet IPv4 par la box ADLS (6rd CE) dotée d'une double pile IPv4/IPv6. Le trafic est acheminé en IPv4 dans le réseau de collecte du fournisseur d'accès (desserte téléphonique, DSLAM...), le tunnel IPv4 est traité par l'équipement frontière entre le réseau de collecte et le réseau de transit de l'opérateur pour être dirigé soit vers le relais 6rd transit/Internet IPv6, soit transit/Internet IPv4 (figure 8.46).

Le 6rd n'utilise pas d'adresse IPv6 spécifique, il utilise les adresses IPv6 du fournisseur qui lui ont été attribuées par son RIR (*Regional Internet Registry*). Le préfixe de l'adresse de la passerelle 6rd CE est obtenu par concaténation du préfixe 6rd attribué par le fournisseur d'accès pour identifier son domaine 6rd et des 24 bits de poids faibles de l'adresse IPv4 attribuée au CE du client complète, il est cependant possible d'utiliser les 32 bits de l'adresse IPv4. La figure 8.47 illustre l'obtention de l'adresse 6rd déléguée à partir du préfixe 6rd « 2001 :0011:/32 » et de l'adresse IPv4 « 10.2.2.3 ».

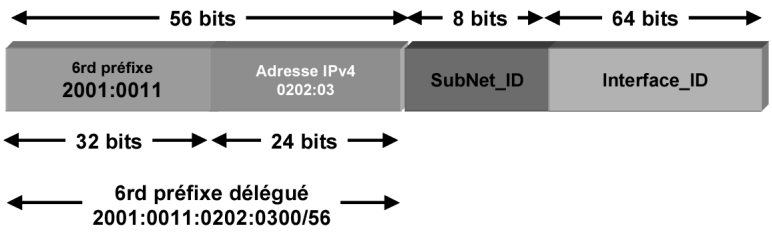

Figure 8.47 Adresse IPv6 d'un client 6rd.

➤ Le NAT64 et DNS64 (RFC 6146)

L'association d'un DNS64[22] à un NAT64 permet à des hôtes IPv6 d'accéder dynamiquement aux services d'un serveur IPv4. L'inverse n'est possible qu'avec une translation statique d'adresses. La figure 8.48 décrit une architecture de type NAT64/DNS64.

Dans cette architecture, le DNS64 fournit l'adresse IPv6 du NAT64 dans laquelle les 4 derniers octets correspondent à l'adresse IPv4 de l'hôte recherché :

1. Le client devant joindre « test.fr » sollicite le DNS64 pour en obtenir l'adresse.
2. Le DNS 64 ne disposant pas de réponse @IPv6 pour cette requête, formule une demande itérative vers un DNS4.
3. Le DNS4 fournit au DNS64 l'adresse IPv4 du serveur recherché.
4. Le DNS64 à partir du préfixe réservé au NAT64 « 64:FF9B::/96 et de l'adresse IPv4 (« 192.26.7.8 » soit « C01A:0708 » en notation IPv6) construit une adresse qui correspond au domaine de réponse du NAT64.

22. La résolution de noms (DNS) est étudiée au chapitre 12 § 2.2

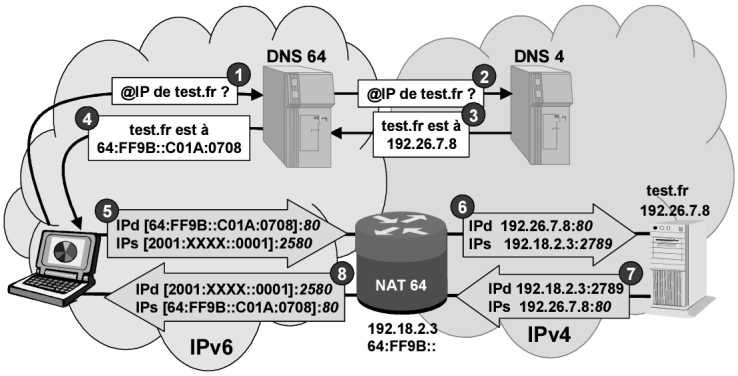

Figure 8.48 Architecture du NAT64/DNS64.

5. L'hôte adresse dons ses messages au NAT64 qui dans ce cas fait office de *proxy* vis-à-vis du serveur test.fr. Notons que le service http est bien appelé au port 80 et que le port source du demandeur est, dans cet exemple, 2580.
6. Le NAT crée une entrée dans sa table et assigne un nouveau numéro de port source[23] (translation de port[24]), et transmet la requête à l'hôte final (test.fr).
7. La réponse de l'hôte n'appelle aucun commentaire.
8. Le NAT64 transfère cette réponse au client en restituant le port source (devenu destination) d'origine.

La translation d'adresses n'est pas la seule opération à réaliser par le NAT. En effet, passant du monde IPv4 et IPv6, le NAT doit gérer la fragmentation notamment en assurant le réassemblage en garantissant l'ordonnancement. Le système est un système à état, celui-ci doit être détruit en fin de session, ce qui est assez commode avec TCP (message « *Fin* »), mais présente une difficulté avec UDP. Dans ce dernier cas, le contexte sera détruit sur temporisation (RFC 4787, valeur par défaut 5 minutes).

➤ Les autres solutions

Les solutions énumérées ici ne sont évidemment pas les seules, à chaque problème à résoudre correspond une solution :

- ❏ Tunnel Broker (RFC 3053) est un système reposant sur des serveurs dédiés gérant automatiquement les demandes de tunnel. Ce système est adapté aux petits sites IPv6 et aux machines isolées sur l'Internet IPv4.
- ❏ Teredo (RFC 4380) est un service qui autorise aux machines masquées par un ou plusieurs NAT d'obtenir une connectivité IPv6.

23. L'association @IP/Numéro de port correspond à l'adresse de transport.

24. Pour plus de détails sur la notion de translation de port (PAT, *Port AdressTranslator*). Cette technique de substitution de ports permet de n'utiliser qu'une seule adresse IP pour connecter en théorie plus de 65 000 machines.

- ISATAP (*Intra-Site Automatic Tunnel Addressing Protocol*, RFC 5214) est utilisé pour le déploiement automatique de IPv6 dans des sites IPv4. ISATAP spécifie le format de datagrammes IPv6 encapuslés dans un datagramme IPv4 (tunnel) dans un réseau IPv4. ISATAP implémente un mécanisme de découverte des routeurs de bordure.
- 6PE (RFC 4798) autorise le transit IPv6 dans un réseau MPLS IPv4. L'implémentation d'une solution 6PE ne concerne que les commutateurs MPLS de bordure, le cœur du réseau n'est pas concerné.
- ...

8.5 CONCLUSION

Défini pour répondre à un besoin précis, TCP/IP n'a été conçu que pour satisfaire les besoins du département de la Défense américain, de ce fait il a bénéficié d'une approche très pragmatique : faire uniquement ce que l'on attendait de lui, c'est-à-dire fédérer des réseaux, les données n'étant que de type texte. Si, a priori, cette approche en limite les possibilités, en fait, elle en a à l'époque facilité les évolutions. Aujourd'hui TCP/IP, bien que conçu pour des applications en mode texte, du fait de sa simplicité originelle et de son ouverture, TCP/IP a su s'adapter pour répondre aux transferts de flux multimédias, voire en particulier à celui de la voix sur IP (chapitre 23).